Analysis of Feature Space for Monitoring Persons with Parkinson's Disease With Application to a Wireless Wearable Sensor System

Shyamal Patel, Konrad Lorincz, Richard Hughes, Nancy Huggins, John H. Growdon, Matt Welsh, Paolo Bonato, *Senior Member, IEEE*

***Abstract*—We present work to develop a wireless wearable sensor system for monitoring patients with Parkinson's disease (PD) in their homes. For monitoring outside the laboratory, a wearable system must not only record data, but also efficiently process data on-board. This manuscript details the analysis of data collected using tethered wearable sensors. Optimal window length for feature extraction and feature ranking were calculated, based on their ability to capture motor fluctuations in persons with PD. Results from this study will be employed to develop a software platform for the wireless system, to efficiently process on-board data.**

I. Introduction

Parkinson's disease (PD) is the most common neurodegenerative disease, affecting about 3% of the population over the age of 65 years [1]. The characteristic motor features include resting tremor, *bradykinesia* (slowness of movement), rigidity, and impaired balance. Complications of medical therapies include *wearing off* and *dyskinesias* [3][4]. A reliable quantitative tool for evaluating motor complications in persons with Parkinson's disease would be valuable both for routine clinical care as well as for trials of novel therapies. In an earlier publication [5] we showed that wearable sensors can be used to closely monitor PD motor fluctuations and predict clinical scores with high accuracy.

In this paper we present the work done to develop software for a novel wireless sensor platform for home monitoring of persons with PD. We also analyze the sensor data collected using wired wearable sensors to evaluate two important factors: (1) optimal window length to extract features from data and (2) ranking of features based on their ability to form tight distinct clusters. Finally, we test the ranked features with a simple linear classifier to assess the impact on accuracy in predicting clinical scores of bradykinesia and dyskinesia. We will eventually use this information to develop software for on board data processing on the wireless units.

Manuscript received April 16, 2007. This work was supported in part by the National Institute of Neurological Disorders and Stroke, National Institutes of Health under the grant #R21NS045401-02.

S. Patel and R. Hughes are with the Department of Physical Medicine and Rehabilitation, Harvard Medical School, Spaulding Rehabilitation Hospital, Boston MA 02114 USA (e-mail: spatel19@partners.org, rhughes1@partners.org).

K. Lorincz and M. Welsh are with the Division of Engineering and Applied Sciences, Harvard University, Cambridge MA 02138 USA (e-mail: konrad@eecs.harvard.edu, mdw@eecs.harvard.edu).

N. Huggins and J.H. Growdon are with Department of Neurology, Harvard Medical School, Massachusetts General Hospital. Boston, MA 02114 USA (e-mail: nhuggins@partners.org, jgrowdon@partners.org).

P. Bonato is with the Department of Physical Medicine and Rehabilitation, Harvard Medical School, Spaulding Rehabilitation Hospital, Boston MA 02114 USA and The Harvard-MIT Division of Health Sciences and Technology, Cambridge MA 02139 (corresponding author; phone: 617-573-2770; fax: 617-573-2769; e-mail: pbonato@partners.org).

II. Wearable Wireless Platform

A. Hardware

The sensor platform we are using is the Intel Digital Health Group's Sensing Health with Intelligence, Modularity, Mobility, and Experimental Reusability (SHIMMER). SHIMMER consists of a TI MSP430 microprocessor; a Chipcon CC2420 IEEE 802.15.4 2.4GHz radio; a MicroSD card slot; a triaxial MEMS accelerometer, the Freescale MMA7260Q; and optionally, a Bluetooth radio which allows streaming of sensor data at high rates, as a radio-agnostic solution.

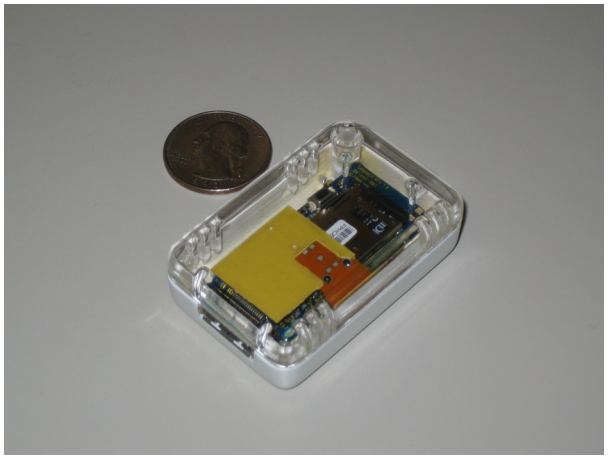

Fig. 1. The SHIMMER wearable sensor platform.

SHIMMER is smaller than other wireless systems, with conventional board technology and a lithium-polymer battery for easy maintenance. Internal and external connectors allow new sensor boards to be interfaced to the device, expanding its capabilities. A triaxial gyroscope board using two InvenSense IDG-300 dual-axis gyroscope chips was designed for internal expansion. One of these gyroscopes is mounted perpendicularly to the main sensor board.

The SHIMMER device combines computation, radio communication, high-fidelity triaxial sensors, and a large flash memory into a tiny, wearable rugged plastic enclosure. It measures 1.75" x 0.8" x 0.5" and weighs just 10 g. SHIMMER utilizes a MicroSD card slot that allows up to 2 GBytes of flash memory. This permits continuous data recording of uncompressed, 50 Hz sampled data from 3 channels for more than 80 days. This amount of on-board storage is unprecedented in wearable sensor design.

B. *Software*

The goal of our software platform is to continuously extract clinically relevant information from the resource-limited nodes worn by patients. Because of the resource limitations in terms of computation, radio bandwidth, and power consumption, we cannot perform all computations in real-time, nor do we have the necessary bandwidth to stream all recorded samples to a base station. Thus we must make several compromises between result delay and result quality. For example, an interesting event may require extensive computations not possible in real time. In this case we can compute some of the data and buffer the rest for later processing. During uninteresting periods (e.g. when a subject is still), we may wish to report only a subset of the features.

Our proposed software architecture consists of four stages. Stage one includes the sampling module, an event detector, and a real-time features module. The sampling component passes samples to the real-time features module. Computed real-time features are then stored to flash and passed to the event detector, which determines if the recorded data is interesting. When needed, raw samples are also stored to flash for further feature extraction. Stage two consists of delayed features modules. Raw samples from storage are read by computationally intensive features algorithms that compute features and store the results back to the flash memory. This stage is scheduled to run opportunistically (i.e. when CPU cycles are available). During stage three, features are transmitted to the base station. This stage also runs opportunistically, only when the patient is within radio range of the base station and spare CPU cycles are available. Finally, in stage four the base station (e.g. a PC or PDA) takes the relevant features from all nodes and passes them through a classifier to predict clinical scores.

III. Data Analysis

A. *Data Collection*

Twelve individuals were recruited in the study, ranging in age from 46 to 75 years, with a diagnosis of idiopathic Parkinson's disease (Hoehn & Yahr stage 2.5 to 3, i.e. mild to moderate bilateral disease) [2].

Figure 2 schematically represents a motor fluctuation cycle. Subjects delayed their first medications intake in the morning so that they could be tested in a "practically-defined OFF" state. This approach is clinically used to observe patients during their most severe motor symptoms. Subsequently, subjects took their medications and were tested every 30-minutes, to gather sensor data across an entire motor fluctuation cycle.

Accelerometer sensors were used to gather biomechanical signals during standardized motor tasks utilized for clinical assessment, including quiet sitting, finger tapping, alternating hand movements, heel tapping, and walking. For each task, 30 s of sensor data were recorded. Accelerometers were placed on the right and left upper arm, right and left forearm, right and left thigh, right shin, and left shin. The sensors were connected to an ambulatory system (Vitaport 3, Temec BV, The Netherlands) to collect and store the signals. Subjects were videotaped throughout the experiment, and motor UPDRS and dyskinesia scores were rated for each task after videotape review

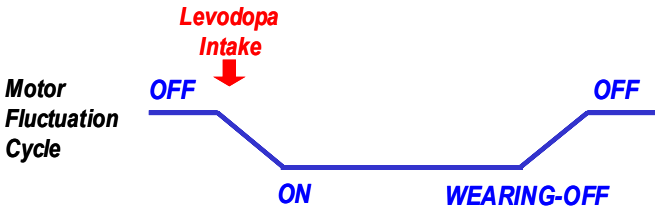

Fig. 2. Schematic representation of a motor fluctuation cycle. See text for details.

B. *Feature Extraction*

Six different types of features were extracted from different body segments. Raw data were high-pass filtered with a cutoff frequency of 1 Hz to remove gross orientation changes, and low-pass filtered with a cutoff frequency of 15 Hz to remove high frequency noise. The features were chosen to represent characteristics such as intensity, modulation, rate, periodicity, and coordination of movement. Intensity was measured as the root-mean-square (RMS) value of the detrended accelerometer signal. The modulation of the output of each sensor was used to represent dynamic characteristics of the tasks, and was calculated as the range of the auto-covariance of each channel. Rate of movement was represented by the dominant frequency component below 10 Hz. Periodicity was measured by computing the ratio of energy in the dominant frequency component to the total energy below 10 Hz. Coordination between body segments on the left and right side and proximal and distal segments was captured in three aspects: magnitude (obtained by calculating the correlation coefficient), delay (estimated as the time lag corresponding to the peak of the cross-correlation function) and similarity (measured by the value of the peak of the cross-correlation function). Approximate entropy was used as a measure of signal complexity.

To study dyskinetic movements, features were derived

from the lower extremity accelerometer channels during a task requiring fine motor control of the upper extremities (e.g. finger tapping). Data features for bradykinesia were extracted from body segments relevant to tested tasks.

C. Optimal Window Length

Subjects performed each task for approximately 30s and features were extracted from 30 randomly selected windows of fixed length. Very small window length will mean that we are not capturing enough information and hence the clusters in the feature space will not be well separated. On the other hand if the window length is too large we will not have enough variance in the cluster to train the classifier on.

D. Feature Ranking

A clustering approach was used to evaluate and rank the data features. Features that form tight and well-separated clusters lead to better predictive performance and vice versa. We chose the Davies-Bouldin (DB) clustering evaluation index [6] because it is well suited to the cloud structure of the feature space.

E. Prediction Performance

The third step was to sequentially evaluate the prediction performance of the ranked features. A simple linear discriminant classifier with a 10-fold cross-validation was used. This will help us to understand how a particular feature affects the prediction performance and also determine if the feature ranks correlates well with prediction performance. We could have used a more complex nonlinear classification technique but since we are only interested in finding the impact of a particular feature on the prediction performance we decided to go with a simple classifier. A more complex classifier will lead only to a better prediction performance.

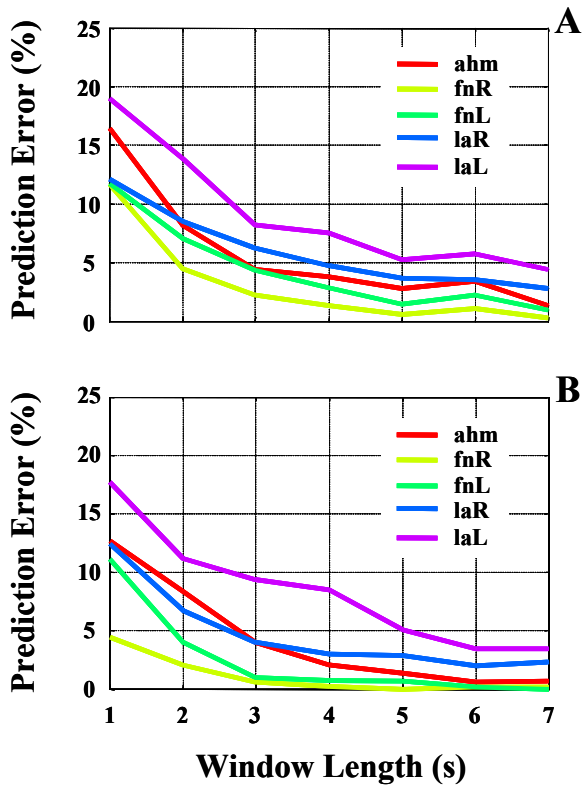

Fig. 3. Plot of prediction error vs. window length (a) Dyskinesia and (b) Bradykinesia. See text for details.

IV. Results

We analyzed dyskinesia and bradykinesia. Tasks were selected for analysis based on their usefulness in capturing motor fluctuation behavior with accelerometers. For dyskinesia we chose alternating hand movements (ahm), finger to nose (fnR/fnL) and finger tapping task (ftR/ftL). For bradykinesia we chose alternating hand movements (ahm), finger to nose (fnR/fnL) and heel tapping (laR/laL).

First we determined optimal window lengths. The window lengths varied in 1s increments from 1 to 7s. A simple linear discriminant classifier was trained with features extracted using different window lengths and tested for prediction performance using a 10-fold cross validation method (Figure 3). A threshold of 3% improvement in prediction error per second of window length was set as the cutoff for selecting optimal window length. For dyskinesia the optimal window length was 5 seconds and for bradykinesia it was 6 seconds.

TABLE I
FEATURE RANKING FOR BRADYKINESIA

%		Features					
		RMS	autocov	corr	maxfrq	maxrat	entropy
Rank	1	32	4	8	8	16	32
	2	24	8	16	4	24	24
	3	12	16	24	4	20	24
	4	12	8	32	20	12	16
	5	20	36	12	28	4	0
	6	0	28	8	36	24	4

TABLE II
FEATURE RANKING FOR DYSKINESIA

%		Features					
		RMS	autocov	corr	maxfrq	maxrat	entropy
Rank	1	20	8	0	4	4	64
	2	68	8	4	0	0	20
	3	8	36	8	20	16	12
	4	4	24	16	36	20	0
	5	0	16	20	28	36	0
	6	0	8	52	12	24	4

Ranking of features was performed based on the clustering index provided by the Davis-Bouldin index for cluster validity. A small value of DB index for a feature means that the clusters are tight and visa versa. Ranked data from 5 subjects was selected because it showed at least a 1-point change on the clinical rating scales for both bradykinesia and dyskinesia. Five tasks for each of the 6

features were evaluated. Tables 1-2 show the percentage time that a feature was given a particular rank by the DB index across all tasks and subjects. Entropy and RMS values emerge as the best features and correlation features are the worst for dyskinesia. For bradykinesia we don't see a clear winner but entropy and RMS value seem to perform a little better than other features. Also the frequency feature (maxrat) performs better for bradykinesia, as it is associated with speed and smoothness of movement.

A linear classifier was constructed with one feature at a time to predict the clinical score. Figure 4 (a) and (b) show that the prediction results correlate well with the output of the feature ranking using the clustering approach. Entropy and RMS value features perform significantly better than other features for predicting dyskinesia scores. For Bradykinesia, features including entropy, RMS value and correlation features seem to.

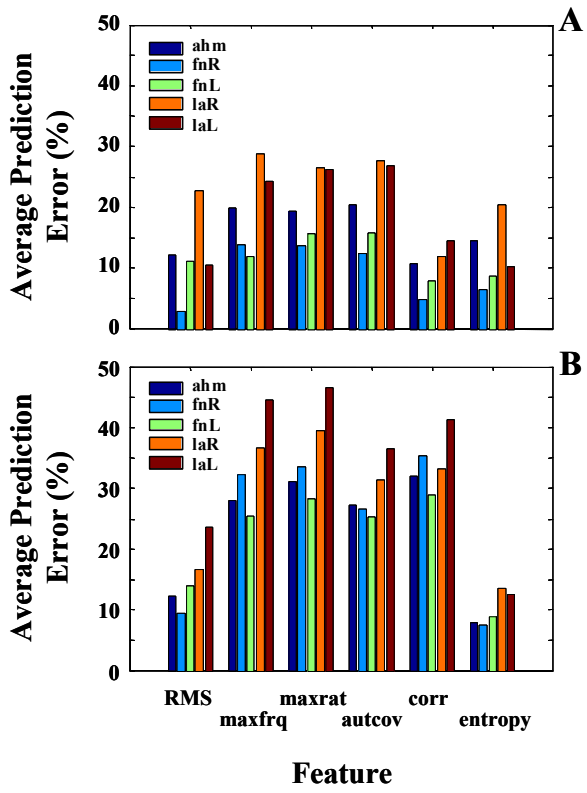

Fig. 4. Prediction performances of individual features (a) Bradykinesia and (b) Dyskinesia.

V. Discussion

The main challenge, in developing a wireless wearable sensor system for home monitoring, is to have a software platform that is accurate yet efficient. With this view we focused on two important factors. We derived an optimal window length for extracting features for two important motor fluctuations of PD viz. dyskinesia and bradykinesia. We ranked the features based on their ability to form well-defined clusters. The ranking we obtained reflects the character of the motor fluctuation i.e. RMS and entropy were dominant for dyskinesia while frequency and correlation features did well along with RMS and entropy for bradykinesia. A combination of two or more of these features will yield good prediction results. Finally, the prediction results obtained using single feature set for the classifier appears to correlate well with the ranking of features. The next step will be to evaluate the computational complexity for doing on-board processing on the wireless sensor platform with the results presented in this paper.

Acknowledgment

The authors would like to thank Dr. David Standaert and Dr. Theresa Lie-Nemeth for their contribution to the study.

References

[1] DG Standaert, AB Young, in Goodman and Gilman's Pharmacological Basis of Therapeutics, Tenth Edition, Hardman JG and Limbird LE, Editors, McGraw-Hill, 2001, pp. 549-620.

[2] S Fahn, RL Elton, "Unified Parkinson's Disease Rating Scale". In Fahn S (Ed), Recent Developments in Parkinson's Disease, MacMillan Healthcare Information, 1987, pp. 153-163.

[3] AE Lang, AM Lozano, "Parkinson's disease. First of two parts", N Engl J Med, vol. 339(16), 1998, pp. 1044-1053.

[4] AE Lang, AM Lozano, "Parkinson's disease. Second of two parts", N Engl J Med, vol. 339(16), 1998, pp. 1130-1143.

[5] Shyamal Patel, Delsey Sherrill, Richard Hughes, Todd Hester, Theresa Lie-Nemeth, Paolo Bonato, David Standaert, Nancy Huggins, "Analysis of the Severity of Dyskinesia in Patients with Parkinson's Disease via Wearable Sensors," International Workshop on Wearable and Implantable Body Sensor Networks, 2006, pp. 123-126.

[6] JC Bezdek, NR Pal "Some new indexes of cluster validity", IEEE Transactions on System, Man, and Cybernetics-part B: Cybernetics, Vol. 28, No. 3, June 1998, pp. 301-315.